第3章

伝送路や開発システムに合った MPEGデータ構成選択の手引き

── 与えられた条件の中でもっとも良い画像を作り出すために

川田亮一

圧縮データを蓄積・伝送するシステムに求められる条件はさまざまである．例えば，テレビ会議では，相手の顔の精細さよりも，伝搬遅延の少ない画像が求められる．逆にコンサートやスポーツ中継では，数秒の遅延があっても高精細な画像が求められるだろう．ここでは，符号化パラメータについて簡単に説明しながら，用途に合った符号化パラメータの選択法を考察する．（編集部）

ひと言でMPEGといっても，その規格上，さまざまな符号化のしかたがあり，このパラメータ選択（データ構成選択）のしかたによって，画質や遅延の程度が変化します．本稿では，さまざまな符号化のニーズに応じて，どのようにデータ構成を選択したらよいのか，そのガイドラインを示します．

現在，MPEG-2やH.264などの動画像圧縮の標準方式が広く使われています（**図1**）．オリンピックやワールド・カップ・サッカーなど，テレビの国際中継は，MPEG-2のCODEC（符号化装置と復号装置）を利用してディジタル圧縮してから伝送するのが普通です．トリノ・オリンピックは，すべてハイビジョンで伝送されました（**図2**）．そのときの圧縮方式はMPEG-2でした．

MPEG-2が標準化されたのは1995年です．それから10年余り経過した現在，MPEG-2に比べると2倍の圧縮効率（圧縮による画質劣化量が同じでも，ビット・レートが半分ですむ）を誇るH.264という新しい標準方式に準拠した

年	規 格
1990	ITU-T　H.261（テレビ電話・会議）
1991	MPEG-1（ビデオCD）
1995	MPEG-2（H.262）（テレビ中継・放送，SDTV/HDTV，広く普及）
1999	MPEG-4（MPEG-2より低レートへ適用用途を広げる）
2003	ITU-T H.264（MPEG-4 AVC） （低レートから高レートまでカバー）

図1　民生機器をはじめとする各種製品に採用された圧縮方式
現在はMPEG-2やH.264などの動画像圧縮の標準方式が広く使われている

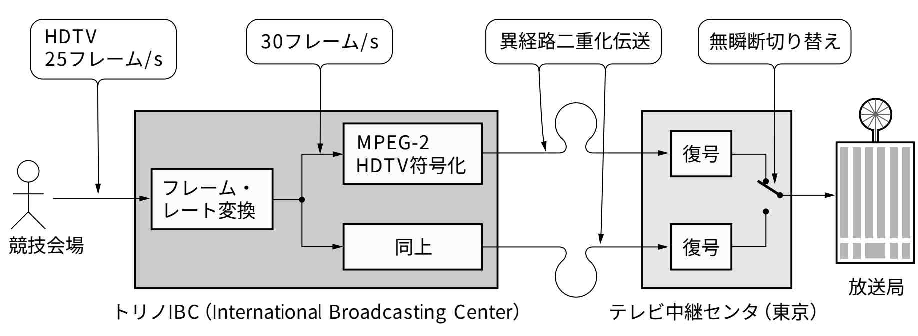


図2
トリノ・オリンピックにおける日本までの映像の伝送経路と主要な機器
トリノ・オリンピックはすべてハイビジョンで伝送された

KeyWord　MPEG-2，H.264，NAB，National Association of Broadcasters，デコーダ，復号装置，エンコーダ，符号化装置，動き補償，フレーム間予測符号化，量子化ステップ，固定レート，可変レート，符号化パラメータ



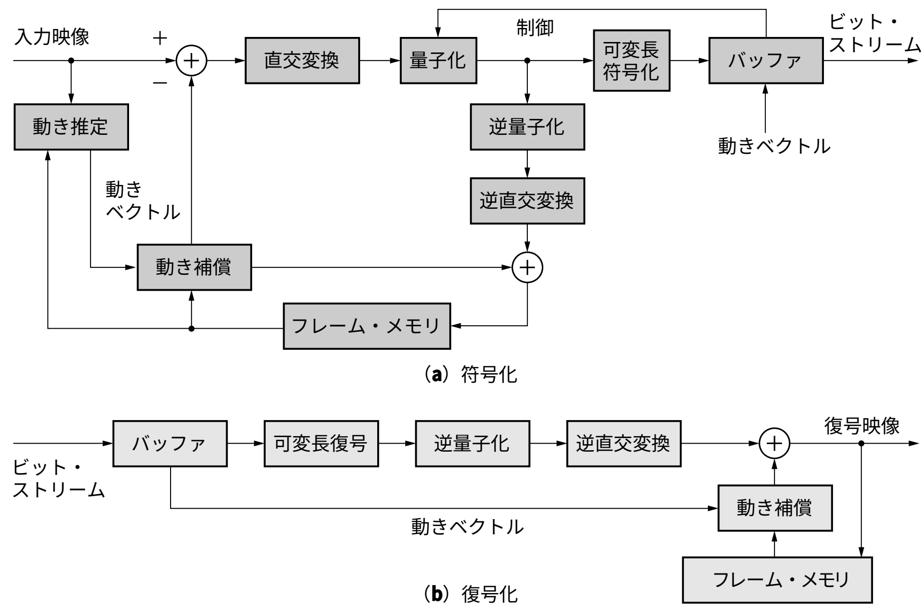


図3
動き補償フレーム間予測符号化の基本構成
MPEG-2もH.264も，動画像符号化のフレームワークとしては，動き補償フレーム間予測符号化方式を共通に用いている

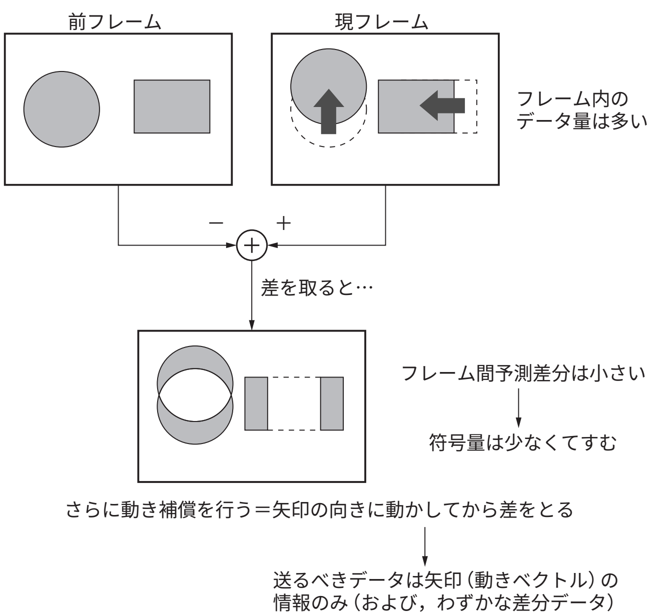


図4　フレーム間予測符号化
直前のフレームからの差分を計算して，その差分だけを圧縮する

CODECが製品化され始めています．2006年4月に開催されたNAB（National Association of Broadcasters）主催のNAB2006（全米放送機器展）でも，H.264のHDTV CODECが各社から発表されていました．

MPEG-2もH.264も，動画像符号化の土台としては，「動き補償フレーム間予測符号化」という方式（**図3**）を共通に用いています．これらは，デコーダ（復号装置）の動作を規定した標準であり，エンコーダ（符号化装置）は，このデコーダで復号できるビット・ストリームを出力するものであれば，すべて標準準拠ということができます．逆に言うと，MPEG-2標準準拠，H.264標準準拠といっても，エンコーダのつくりや符号化時のパラメータ選択によって，復号画質や伝送遅延量にはいろいろと違いが出てきます．

デコーダは，基本的には送られてきたビット・ストリームを復号するだけであり，画質はすでにエンコーダ側で決定されているといえます．そこで以下では，

- 画質を維持しながらフレーム・レートを落とす
- フレーム・レートを落とさずにビット・レートを落とす
- 送り側と受け側の遅延があってもよいからビット・レートを落としたい
- リアルタイム性を重視する
- ネットワークで転送する

などの観点から，符号化パラメータ決定（データ構成選択）のガイドラインを示します．

● 伝送ビット・レートの大小は予測の精度に依存する

まず準備として，MPEG-2やH.264の基本的な枠組みである動き補償フレーム間予測符号化について説明しておきます．ハイビジョンも含めてテレビの映像信号は，1秒間に30枚のコマ（フレーム）からなっており，これを順次圧縮していきます．その際なるべく情報量を少なくするために，各フレームごとに独立に圧縮するよりも，直前のフレームからの差分を計算して，その差分だけを圧縮するとい

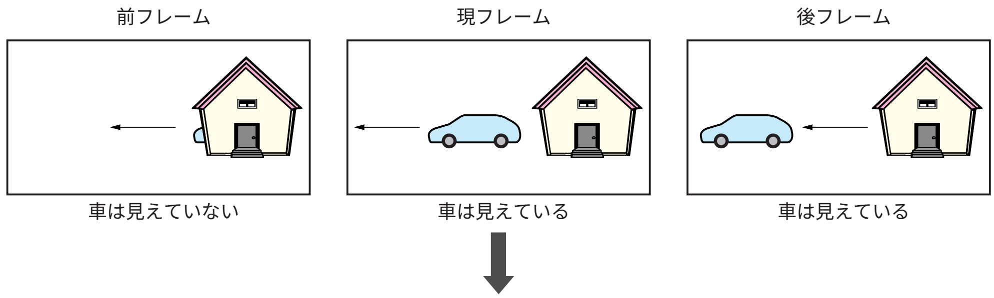


図5
平均化画像による予測
直前・直後だけではなく，もっと離れたフレームからの予測も可能

う方法が有効です．これをフレーム間予測符号化といいます(**図4**)．

さらに動画像中にはさまざまな動きがあるため，物体ごとにその動きを追ったうえで差分を計算すれば，単純に画面全体の差分をとるよりも，符号化すべき情報量を小さくできます(各部分ごとの動きの情報は付随情報として伝送する)．これを動き補償フレーム間予測符号化方式といいます．

この動き補償予測効率を向上させる(符号化すべき情報量を小さくする)ために，MPEG-2やH.264では，直前のフレームによる予測だけではなく，直後のフレームと直前のフレームの平均化画像による予測も行える規則になっています(**図5**)．

前のフレームでは物体の陰に隠れていて見えなかった部分がある場合には，後のフレーム(その部分が見えている)を使ったほうが，差分を小さくできるのです．さらに直前・直後だけでなく，もっと離れたフレームからの予測もできます．これらは上述の符号化パラメータ選択時の選択肢となります．

ただし，時間的に後のフレームを「予測」に使うということは，符号化の順番を，実際のフレームの順番から変える必要が出てきます．何フレームか後の画像を先に復号してからでないと，現フレームを復号したり表示したりできません．これは伝送処理遅延の増大につながります(**図6**)．

一方，時間的に前のフレームだけが予測に使用されていれば，そのような遅延は発生しません．この処理遅延量の大小も，パラメータ選択の際に考慮すべき点です．

● 画質は量子化ステップ・サイズで決まる

また，固定レートと可変レートという選択肢もあります．もとの映像の情報量はつねに一定ではありません．精細な絵柄で複雑な動きを含むシーンでは情報量は大きく，単純な絵柄や静止画像の場合には，情報量は小さくなります．このようにシーンごとに異なる情報量を持つ映像を圧縮するには，圧縮後の情報量も変化してよい(可変レート)ということを前提とするほうが，符号化制御としては簡単です．つまり，予測差分信号を一定の値(量子化ステップ・サイズ)で割り算して，その商を伝送する，固定量子化という方法です．

ところで，固定かどうかにかかわらず，量子化のステップ・サイズが大きいほど，伝送するべき商の値は小さくなるため，情報量を削減できます．その代わり，復号側で量子化ステップ・サイズと商を掛け算してもとの値を推定する際の誤差が大きくなる，つまり画質劣化が大きくなります(**図7**)．

量子化ステップ・サイズによって画質が決定されるので，これが固定ということは，画質が一定ということになります．その代わり，大きな情報量を含むシーンでは，もとの予測差分信号も大きいため，圧縮後の発生情報量も大きくなります．つまり可変レートとなります．

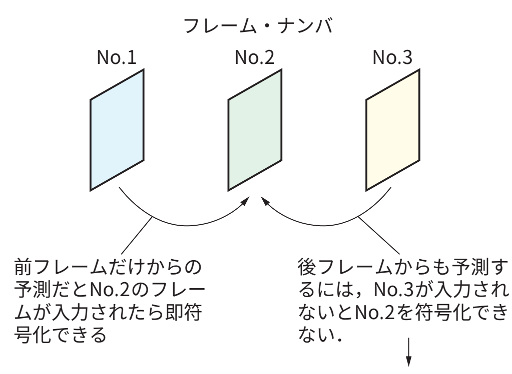


図6　エンコードで遅延が生じる理由
処理遅延量の大小もパラメータ選択の際に考慮するべきである

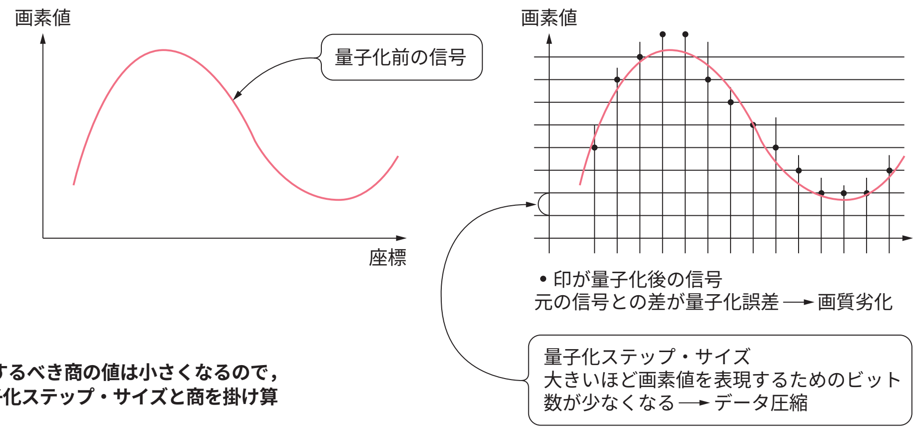


図7
量子化で画質が劣化する理由
量子化のステップ・サイズが大きいほど，伝送するべき商の値は小さくなるので，情報量を削減できる．その代わり，復号側で量子化ステップ・サイズと商を掛け算して元の値を推定する際の誤差が大きくなる

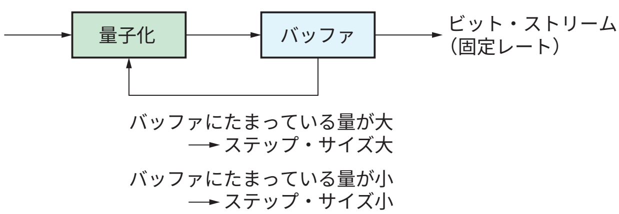


図8　可変量子化
固定レートでは，発生する情報量がなるべく一定となるように，量子化ステップ・サイズを制御する

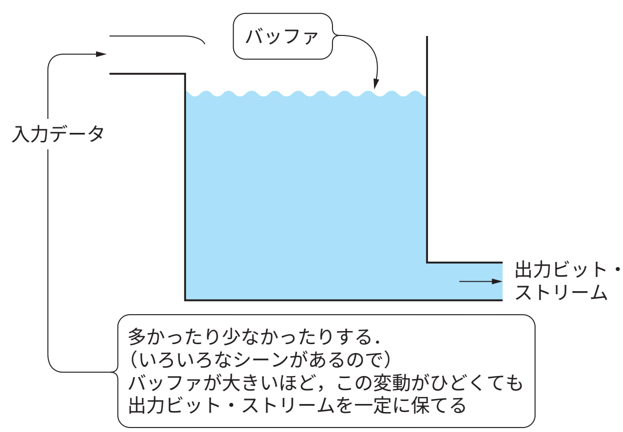


図9　ダムの貯水量が大きいと遅延も大きくなる
バッファが小さいと，画質的には不利．元の映像の情報量が小さい場合に，バッファにたまった符号量がたまたま十分でないと，伝送ビット・レートを一定に保つ目的でダミーのビットを追加する必要が生じる

可変レートは，テープやディスクに映像を記録する場合にはよいのですが，テレビ中継などの伝送用としては，伝送路のビット・レートの上限が決まっているため，そのままでは不適切です．ここで固定レートという考えかたが出てきます．

固定レートでは，発生する情報量がなるべく一定となるように，量子化ステップ・サイズを制御します．つまり，元の情報量の多いシーンでは量子化ステップ・サイズを大きく（粗く）し，情報量の少ないシーンでは小さく（細かく）します（**図8**）．

ここで「発生する情報量を一定にする」と言っても，どの単位で一定にするのかということが課題となります．つまり，1フレームごと（1/30秒）なのか，1秒間（30フレーム）ごとに一定になればよいのか，というようなことです．

いずれにせよ伝送路のビット・レートは決まっているので，後者の場合，1秒間分の符号化データをエンコーダの出口手前でダムのように蓄えるバッファが必要です．

ダムからの放水量は一定（発生情報量一定）ですが，その分，ダムの貯水量はシーンに応じて上下してもよいのです．ただしバッファの分の処理遅延が生じます．

一般に，ダムの貯水量（バッファの大きさ）が大きいほうが，符号化制御は簡単になります．ただしため込む量が大きくなる分，処理遅延も大きくなります（**図9**）．しかしバッファが小さい（処理遅延が小さい）と，画質的には不利となります．元の映像の情報量が小さい場合に，バッファにたまった符号量がたまたま十分でないと，伝送ビット・レートを一定に保つためにダミーのビットを追加してやる必要が出てきます．つまり，むだなビットを送る必要があります．バッファが大きければ，シーンによる情報量の大小を平滑化でき，このような伝送ビットのむだが生じにくくなるのです．

以上のように画質[注1]は主として量子化ステップ・サイズによって決まり，そのときの伝送ビット・レートの大小は，

注1：ここでいう画質とは，「圧縮していない画像」と「符号化し，復号処理を行ったあとの画像」の違いを指す．

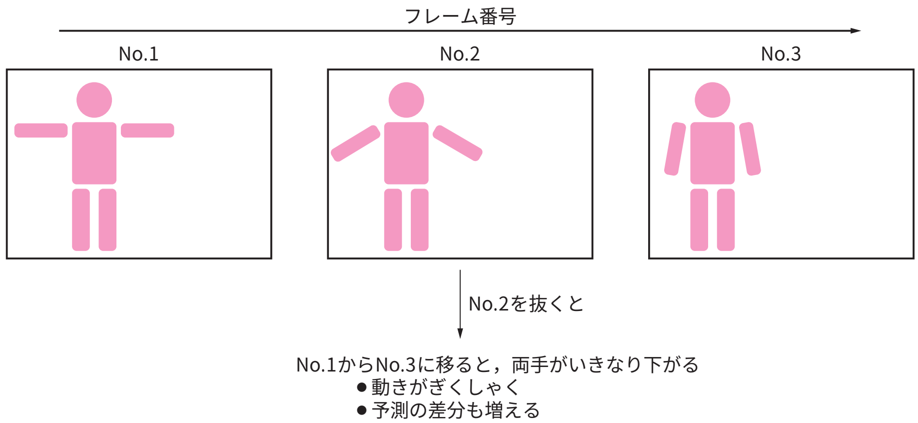


図10
フレーム・レートを落とすとビット・レートは落とせるが，動きそのものはぎくしゃくする
フレームとフレームの時間間隔が広がるため，予測は当たりにくくなる

予測がよく当たっているかどうかに依存することになります．

● 画質を維持しながらフレーム・レートを落とす

フレーム・レートを落とすと，フレームとフレームの時間間隔が広がるため，予測は当たりにくくなります(図10)．しかし，符号化対象が少なくなるのですから，一般に同じ画質でも伝送レートが小さくてすみます．

ただし，ここでいう画質とはフレーム1枚1枚の画質であり，動きの滑らかさという意味では，フレーム・レートを落とすと，当然，不利になります．

そこで，フレーム・レートを落とすという行為は，動きの滑らかさを犠牲にしてでも，1枚1枚のフレームの画質を優先したい場合に適していると言えます．例えば，監視カメラの映像などです．囲碁や将棋の映像を送ったりする用途などにも向くかもしれません．

● フレーム・レートを落とさずにビット・レートを落とす

フレーム・レートを落とさずにビット・レートを落としたい場合，1枚1枚の画質を落とすしかないので，量子化ステップ・サイズを大きくすることになります．ただし，伝送遅延を許容するのなら，なるべく予測を強力(高度)なものにして，差分信号を小さくします(図11)．

放送局におけるテレビ中継においては，一般の放送と比べてはるかに高い画質の伝送を行う必要があります．この場合，フレーム・レートを落とすことはせず，SDTV方式で15Mbps程度，HDTV方式で30Mbps程度(MPEG-2圧縮)のビット・レートを要します．

● 遅延があってもよいからビット・レートを落としたい

送り側と受け側の遅延を許容するなら，予測を高度にして，なるべく多くのフレームから予測できるようにします．究極的には可変レート(固定量子化)が，同じビット・レートならもっとも画質は良くなります．ニュース番組などで，アナウンサと現地の記者がやり取りするような場合は，なるべく遅延が少ないほうが望ましいのですが，番組を一方的に配信・放送する用途では，遅延は大きくてもそれほど問題にならないので，画質のチューニングを優先できます．

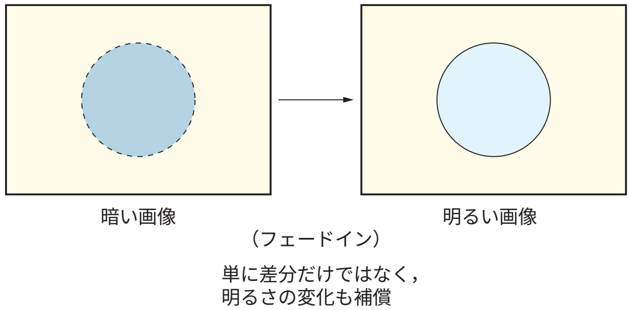


図11　高度な予測の例(明るさまで補償する予測)
伝送遅延を許容するのなら，なるべく予測を強力(高度)なものにして，差分信号を小さくする

また，ハード・ディスクなどにいったん蓄える場合も，遅延は関係ありませんから，可変レートでじっくり符号化できます．

● リアルタイム性を重視する

双方向の会話が含まれる場合には，遅延はなるべく少ないに越したことはありません．もっとも遅延を小さくしたいときには，フレーム間予測を使用しないで，フレームごとに一定のビット・レートで符号化します．

つまり，フレームによって，量子化ステップ・サイズが異なってきます．このとき，各フレームの中では，なるべ

3


（a）発生ビット量が多い　（b）発生ビット量が少ない

写真1　発生ビット量が多い画像と少ない画像の例

く同一のステップ・サイズとなるように制御したほうが，画質は良くなります．

しかし，このフレーム内のステップ・サイズの制御は，符号化装置内部の制御アルゴリズムに関する部分であり，普通，回路設計者は触れることはできません．最近では，この遅延を避けるために，非圧縮でHDTV放送を光ファイバ伝送するケースもあります．

● ネットワークで転送する

IPネットワーク上の映像ストリーミングの場合，可変レートでよければ，上記のようになるべくフレーム間で量子化ステップ・サイズをそろえたほうが，画質が良くなります．そうは言っても，ピーク・レートは抑えたほうがよい場合が多いでしょうから，あまり発生ビットが多くなるときは（**写真1（a）**），ステップ・サイズも小さくするという制御が適しています．つまりピーク・レートを規定したうえで，それ以下については可変レートとします．

単にいったんファイルとして蓄積した映像を転送するだけなら，通常のFTP（file transfer protocol）などが使えるため，とくに問題はありません．ただし，ファイル・サイズが大きい場合には，途中で止まってしまったりした際の対策（スムーズな再開など）を施した特別なツールを使ったほうがよいこともあります．

かわだ・りょういち
（株）KDDI研究所

<筆者プロフィール>
川田亮一．1991年，東京大学大学院工学系研究科 修士課程修了．同年，KDD（現KDDI）入社，研究所に勤務．テレビ中継に関する研究開発，とくに可変長符号化復号LSI，動き補正TV方式変換，次世代映像符号化，高能率無瞬断映像伝送，映像画質監視，次世代多視点映像伝送などの研究に従事．1997-1998年，米コロンビア大学 客員研究員．2001年，映像情報メディア学会技術振興賞開発賞，研究奨励賞，鈴木記念賞受賞．2006年よりKDDI技術戦略部勤務．

ディジタル信号処理シリーズ　好評発売中
ディジタル画像の基礎からJPEG2000・電子透かしまで
よくわかる動画・静止画の処理技術
貴家仁志 著　B5変型判 196ページ　CD-ROM付き　定価2,940円(税込)　ISBN4-7898-3399-2
CQ出版社　〒170-8461 東京都豊島区巣鴨1-14-2　販売部 TEL.03-5395-2141　振替 00100-7-10665